VISUAL COGNITION, 2006, 13 (5), 623–642

Psychology Press
Taylor & Francis Group

The role of attention in processing configural and shape information in 3-D novel objects

Simone K. Favelle and Stephen Palmisano
Department of Psychology, University of Wollongong, NSW, Australia

Darren Burke
Department of Psychology, Macquarie University, Sydney, Australia

William G. Hayward
Department of Psychology, Chinese University of Hong Kong, Shatin, Hong Kong

Recent research suggests that there is an advantage for processing configural information in scenes and objects. The purpose of this study was to investigate the extent to which attention may account for this configural advantage. In Experiment 1, we found that cueing the location of change in single object displays improved detection performance for both configural and shape changes, yet cueing attention away from the location of change was detrimental only for shape change detection. A configural advantage was present for each cueing condition. Experiments 2A and 2B examined whether the configural advantage persisted in conditions where attention was distributed more widely, using a visual search paradigm. Although searches for configural changes were faster than those for shape changes across all set sizes, both types of information appeared to be processed with similar efficiency. Overall, these results suggest that the configural advantage is independent of the location or distribution of visual attention.

A central question in visual cognition and perception is how we process complex entities such as 3-D objects. The objects we perceive in our visual environment possess many different properties including colour, texture, size, orientation, and motion. However, a common approach to the issue of visual object processing is to assume that the visual system represents objects in terms of parts; in

Please address all correspondence to: S. Favelle, Department of Psychology, University of Wollongong, Northfields Ave., Wollongong 2522, NSW, Australia.
Email: simone_favelle@uow.edu.au

We would like to thank Glyn Humphries and two anonymous reviewers for their helpful comments and suggestions in relation to earlier drafts of this manuscript.

http://www.psypress.com/viscog DOI:10.1080/13506280544000228

particular, their shape and the spatial relations between them (e.g., Biederman, 1987). With this assumption comes the issue of the way in which parts and their configural relations are processed and integrated. For the purpose of the current discussion, we will define: (a) a ''part'' as a restricted portion of the object that has semiautonomous, object-like status in visual perception (Palmer, 1999); and (b) a ''configuration'' as the locations in space occupied by these parts. As with Pomerantz's (1983) place relationships, configuration is not dependent on the identity of the parts.

Previous research has investigated the processing of parts and their relations during tasks involving 2-D object recognition. Studies have consistently demonstrated the primacy of configural over part component information in both common and novel 2-D object processing (Cave & Kosslyn, 1993; Kimchi & Bloch, 1998). More recently, this research has been extended to the study of 3-D objects. Specifically, Keane, Hayward, and Burke (2003) investigated the configural and part shape information used in detecting changes to 3-D novel objects. Three object properties were examined: The configuration of an object's parts, the shape of those parts, and their relative arrangement. Using a one-shot change detection task, Keane et al. found that observers detected changes to the object's overall configuration quicker and more accurately than changes to the shape of the parts or a switching (changing the arrangement) of parts. These differences in change detection performance for the three object properties remained when a quantitative measure of pixel change was analysed as a covariate, suggesting that information regarding the global configuration of parts is better encoded than more local details, such as part shape. These results were similar to those found for change detection involving displays of object arrays (Simons, 1996). Also using a one-shot change detection task, Simons found that detection of changes to the configuration of objects in an array were detected more accurately than changes involving two of the objects switching locations or changes to the identity of the objects. Together these studies suggest that there is a configural advantage in detecting change to visual object information.

The main purpose of the current study is to investigate the extent to which attention may account for this configural advantage. Any change that occurs in our visual field is typically accompanied by a motion transient signal that attracts our attention to the location of the change (e.g., Klein, Kingstone, & Pontefract, 1992). When this signal is masked, however, detecting change becomes quite difficult (see Rensink, 2002; Simons & Levin, 1997, for reviews). As a consequence, attention has played a key role in many previous explanations of change detection findings. For example, Rensink, O'Regan, and Clark (1997) have argued that change detection results can be explained by the idea that focused attention is necessary for detecting change. They demonstrated that attention to the relevant portion of a scene increases the likelihood of successful change detection. In their study changes were made to an object or area of central or marginal interest, as rated by independent observers, with central interest areas assumed to attract more attention than marginal interest areas. Rensink et al. found that changes to central interest objects were detected more accurately than changes to marginal interest objects. Thus, they argued that focused attention is necessary for storing scene elements in memory and therefore necessary for change detection.

However, attention to the location of change is not always sufficient for successful change detection. This notion is supported by studies showing change blindness for central actors in video and real-world sequences (Levin & Simons, 1997; Simons & Levin, 1997). Fernandez-Duque and Thornton (2000) demonstrated that even when subjects were unaware that a change had occurred, they could reliably select the changed item in a forced choice task. Collectively, these studies suggest that rather than the location of change; certain other aspects of a scene may be attended.

One possibility is that observers might attend to the global layout or form of scene elements, as this provides a framework or backdrop within which other items can be placed. O'Regan, Deubel, Clark, and Rensink (2000) provided support for this notion with a study that analysed the eye movements of observers during scene change detection tasks. They found that the probability of detecting a change was greater when the location of change was closer to the eye's fixation. However, even when observers directly fixated the location of the change (within 1 degree of visual angle), they still failed to detect the changes a large proportion of the time. O'Regan et al. argued that what the observer sees in a scene is not necessarily determined by which location is being fixated, but rather by which scene aspects are being attended. A scene aspect may consist of a grouping of a subset of scene elements (e.g., the configuration of elements) that may set up a framework for other items in the scene.

Another task used to investigate the role of attention in object and scene processing is visual search. Similar to change detection, visual search is based on detecting differences between stimuli. However, in a visual search task, a subject looks for a designated target item amongst a number of irrelevant or distracting items. Thus, the visual search paradigm is often used as an analogue of a more realistic visual situation such as searching for a friend in a crowd. This task allows researchers to examine how objects are differentiated, which stimulus properties attract attention, how attention is deployed from one object to the next, how one keeps track of what is attended, and how efficiently stimulus properties can be processed (Chun & Wolfe, 2000). Specifically, visual search efficiency (processing time per item in the display) is typically assumed to reflect the ease with which target items are discriminated from distractors.

Saumier, Arguin, Lefebvre, and Lassonde (2002) used a visual search task to examine the way in which a visual agnosic patient (AR) encoded object parts and the relations between those parts. The stimuli used were three different types of 3-D objects (resembling either a four-legged animal, a bird, or a plug) constructed from basic volumetric geon-like parts. There were four target types: (a) Those with the same configuration and parts as distractors, (b) those with the same configuration and all different parts to distractors, (c) those with different configurations and same parts as distractors, and (d) those with a different configuration and all different parts to distractors. For all subjects (AR and controls), Saumier et al. found that search rates were faster when the target differed from the distractors in terms of the configuration of parts (compared to when they had the same configuration). That is, search was slowed significantly when the configuration of parts was the same for both the target and distractors. Search was worst when targets and distractors shared both their parts and configuration. This difference was greatly exaggerated in AR's results, such that AR had extreme difficulty in discriminating objects that shared their configuration, regardless of whether the parts were the same or different (slope magnitudes were up to 290% larger for AR than controls for these conditions). These results suggest that different processes are involved in the perception of parts and overall configuration.

Thus, to summarize, implicit in the change detection paradigm is the idea that observers are likely to detect those visual property changes that they are attending. One question we wished to address in the current study was whether this configural advantage might be explained in terms of an attentional bias or preference towards configural properties. On the one hand, it was possible that configural information benefits from a perceptual processing advantage that is independent of attention. However, it was also possible that attention is allocated more easily to the global form of an object (the configuration of its parts) than more local details such as the shape of parts.

Specifically, we investigated: (a) Whether drawing attention to the exact locus of the change would increase the likelihood of successful change detection for novel 3-D objects, (b) whether there would be an interaction between the locus of attention and the processing of object properties (such that focused attention to a location involved in a change improves part change detection, but has little effect on the detection of configuration changes), and (c) whether the configural advantage for object processing would persist in visual search tasks where attention is distributed more widely as the set size increases (as opposed to the narrow distribution of attention possible when viewing single object displays). In Experiment 1, participants were presented with displays consisting of single novel 3-D objects and attention was drawn to a particular part by having that part change in both colour and texture. Using two different visual search tasks in Experiments 2A and 2B, we investigated the processing and detection of visual properties of novel 3-D objects in multiple item displays. In particular, we explored whether the configuration of object parts would be better encoded and processed only in single object displays, or whether the configuration of object parts would be attended regardless of the number of items in a display.

EXPERIMENT 1

Scholl (2000) distinguished between endogenous control and exogenous capture of attention in his investigation of change blindness. Endogenous control of attention requires voluntary direction of attention to an object or location whereas exogenous control of attention is the involuntary capture of attention by some salient aspect of a scene. Previous studies that have investigated attention in change detection have typically used ratings of the ''centre of interest'' to determine whether attention is likely to be directed towards a particular area of an image (O'Regan, Rensink, & Clark, 1999; Rensink et al., 1997). However, these ''centre of interest'' ratings are not ideal measures of attention as they depend on the subject being able to verbalize the change (Scholl, 2000). Using a flicker paradigm, Scholl investigated whether change blindness for common objects was attenuated by exogenous or externally based attentional capture. The changes to be detected were either a replacement change or a flip change to an object within an array of 12 randomly positioned objects. Exogenous capture of attention was produced using either a late-onset item (the appearance of a stimulus where none was before) or colour singletons (the presence of a unique colour in a display, such as a green item in a field of white items, e.g., Theeuwes, 1991, 1992; Yantis & Jonides, 1984). While these exogenous capture manipulations were never reliable cues to the location of change (changes could occur anywhere in the array), Scholl found that change blindness was indeed attenuated when the changed item was late-onset or a colour singleton. This result suggests that changes to these items were detected faster because the changed object was being attended. The results support the attention-based theory of change blindness, that is, the detection of change requires attention.

A similar idea to exogenous capture of attention was used in the current experiment to draw attention to specific parts of a single 3-D object. As in Keane et al. (2003), changes to the configuration of parts and changes to the shape of parts were investigated. The aim of this experiment was to determine whether drawing attention to or cueing the location of a part involved in a change improved detection performance compared to conditions in which no cue was available or when the cue was not at the location of the change. This cueing was done via a colour and texture change; shortly after an object was displayed, the colour and texture of one of the parts changed (see Figure 1).

One potential explanation of previous change detection results for novel objects (e.g., Keane et al., 2003) is that changes to the shape of parts were detected poorly because attention was directed more easily to configural than part defined properties. If this was the case then a valid cue to the location of change should improve performance in detecting changes to part shape. If attention is necessary for shape change detection, then a nonvalid cue to the location of a shape change should hinder performance. If changes to the object's configuration are detected quicker and more accurately than other change types

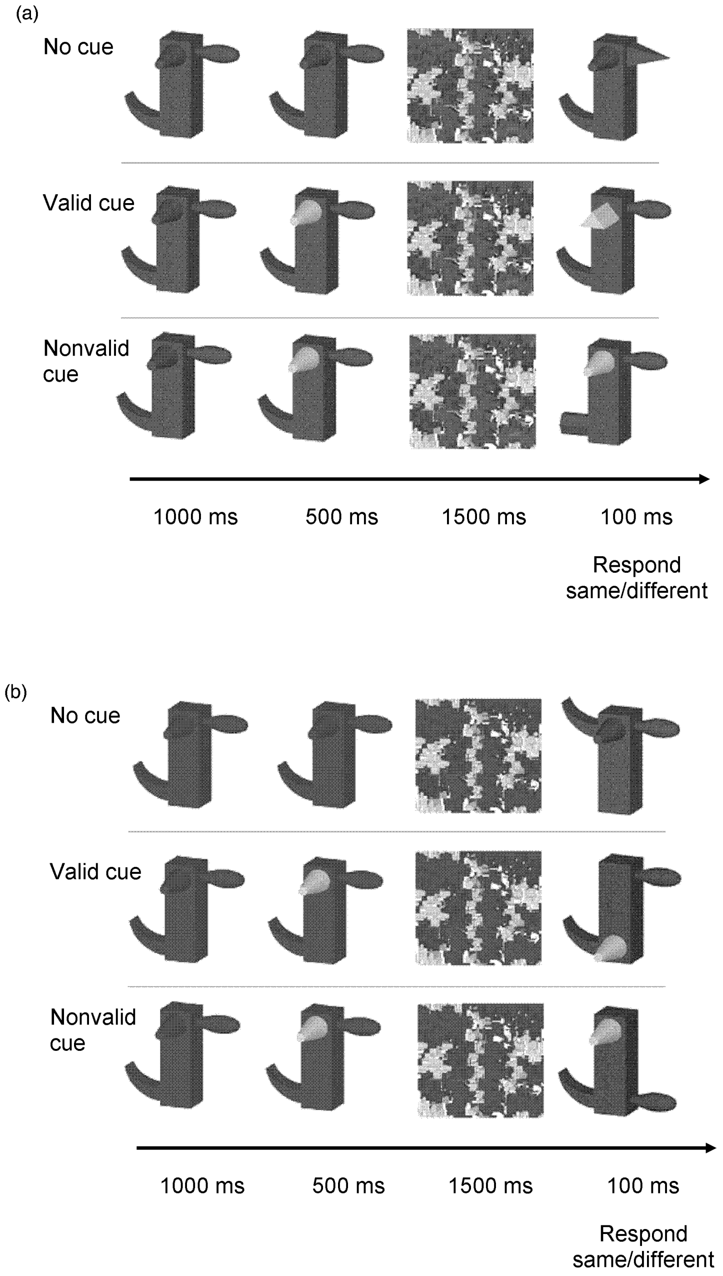

Figure 1a. The three different part (a) shape and (b) configuration change trial sequences involving: No cue to the location of change; a valid colour cue to the location of change; and a nonvalid colour cue to location of change.

because attention is focused on this information more easily than other aspects of the object, then the effect of the attentional cues should be minimal, if any.

Method

Subjects. A total of 20 subjects (18 undergraduate students and 2 academic staff) participated and were tested individually. Undergraduate students received course credit for participating.

Materials. Stimuli were 3-D novel objects. There were three ''standard'' objects. Each object was composed of a main body with three appendage parts. These parts were attached to the body at three of six possible positions. Manipulations to parts were made in terms of configuration or shape. Each object was rendered three times: (a) In a single colour, (b) in two colours such that one part involved in a change was a different colour to the rest of the object, and (c) in two colours such that one part not involved in a change was a different colour to the rest of the object (see Figure 1). This gave a total of 93 different object exemplars used in the current experiment. All objects were photorealistically rendered with the same colours and textures. The entire background screen was white. The mask was 500 × 400 pixels in area. All experiments reported in this paper were controlled by RSVP software (Williams & Tarr, 1998) on Macintosh computers with 15-inch Macintosh CRT (640 × 480 pixels).

Procedure. Each trial began with a fixation cross appearing for 0.5 s at the centre of the screen, followed by the first object for 1.5 s, followed by a mask appearing on the screen for 1.5 s, and finally a second object displayed for 100 ms. Responses for each trial timed out after 5 s. The next trial began 1 s after the subject made a response or the trial timed out. The first object was either all one colour for the 1.5 s or one of the parts changed colour 1 s after stimulus onset. That is, an all blue object was on display for 1 s, then for the remaining 0.5 s either: (a) An all blue object remained on display, or (b) one of the object parts changed colour to green. When one of the parts of the first object changed colour it was either a valid or nonvalid cue to the location of change. If the coloured part was a valid cue to the location of change, it was involved in either a change to part configuration or to part shape. If the coloured part was a nonvalid cue to the location of change, the part was not involved in an object property change (see Figure 1).

The first object in each trial was placed in the centre of the screen; the second object in each trial was randomly placed at a position 25 pixels in any direction from the centre of the screen. Subjects were told that one of the parts of the first object might change colour. Regardless of the colour, subjects were asked to indicate whether the first and second objects were the ''same'' or ''different'' by pressing corresponding keys on a keyboard. Half of the trials were ''same'' trials and the other half ‘‘different’’ trials. The different trials were split equally into the two change type conditions. Feedback was given in the form of a beep to incorrect responses.

Results and discussion

A 3 × 2 repeated-measures ANOVA including cue type (valid, nonvalid, and none) and change type (configuration and shape) was used to analyse accuracy data. A significant main effect was found for change type, $F(1, 19) = 68.54$, $p < .01$. Comparisons of the mean proportion correct (in parentheses) for each condition revealed that configuration changes (.88) were detected more accurately than shape changes (.69). There was also a significant main effect of cue type, $F(2, 38) = 18.69$, $p < .01$. Bonferroni-adjusted post hoc contrasts showed that change detection accuracy for trials with valid cues was significantly greater than that found for either nonvalid cued trials or trials with no cue (both *p*s < .01). Further, the accuracy performance in trials with nonvalid cue trials was significantly poorer than in trials with no cue ($p < .01$). A significant interaction was found between cue type and change type for accuracy, $F(2, 38) = 5.09$, $p < .05$ (see Figure 2). Bonferroni-adjusted post hoc contrasts showed that change detection accuracy for validly cued shape changes was greater than for no cue shape changes and that accuracy for nonvalid cued shape changes was poorer than for no cue shape changes (both *p*s < .01). While validly cued configural changes were detected with greater accuracy than no cue configural changes ($p < .01$), there was no difference in detection accuracy between configural changes that were nonvalidly cued or not cued at all ($p = .52$). Further, shape change detection accuracy with a valid cue was not significantly different to accuracy for configural changes that were either nonvalidly cued or not cued at all ($p = .86$ and $p = .41$, respectively).

Data analysis of RT was conducted using accurate responses. A 3 × 2 repeated-measures ANOVA including cue type (valid, nonvalid, and none) and change type (configuration and shape) was used to analyse RT data (see Figure 2). No main effect of cue type was found for RT, $F(2, 38) = 2.46$, $p = .10$. However, a significant main effect was found for change type, $F(1, 19) = 11.63$, $p < .01$. A comparison of mean RT (in parentheses) for each condition revealed that configuration changes (922.6 ms) were detected quicker than shape changes (964.4 ms). No interaction was found between cue type and change type for RT, $F(2, 38) = 0.98$, $p = .39$.

Relative to the no cue condition, valid cues to the location of change within an object facilitated both configural and shape change, whereas nonvalid cues to the location of change only inhibited detection performance for shape change. With a valid cue to the location of change, shape change detection accuracy could only be improved to the level of configural change with no cue or a nonvalid cue. Detection of configural change was equally successful regardless

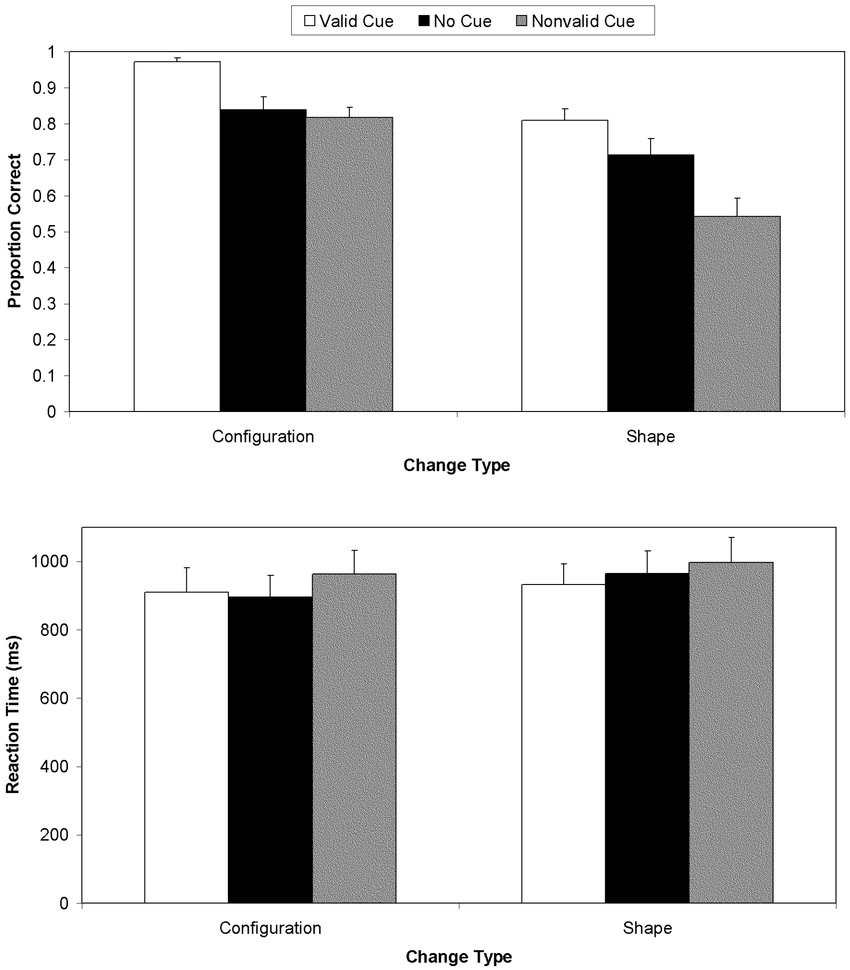

Figure 2. Mean proportion correct (top) and mean reaction time (bottom) on the change detection task as a function of change type and cue. Error bars represent standard errors of the mean.

of whether attention was drawn away from the location of configural change with a nonvalid cue or no cue to change location was provided. This suggests that the facilitatory effect of the valid cue for configural change relative to both of these conditions might be based on additional local change signals, which were indirectly generated by the change in configural information. During a validly cued configural change, two events take place: (a) The part disappears from the attended location, and (b) it moves to a new location on the object. We propose that the disappearance of the part from an attended location may be processed as both a change to local part information, and a change to overall configuration (whereas the disappearance of a part in an unattended location will not be processed as a change to local information, only as a change to configural information). This is consistent with the idea that part-based attention can operate concurrently with the processing of an object as a coherent whole (Vecera, Behrmann, & McGoldrick, 2000). Based on this explanation, we could conclude that the processing of configural information for change detection is unaffected by the locus of attention and that the facilitatory cueing effect is based on additional part change information (which supplements the configural change information).

The only significant effect in the RT data was that configural changes were detected faster than shape changes. Thus, while attention cued to the locus of change improved the accuracy of detection performance, it had little effect on the time taken to process different object properties. Overall, the implications of these results are that: (a) Focused attention is necessary for the accurate detection of changes to shape properties; (b) while focused attention can improve the accuracy of configural change detection, it is not requisite; and (c) irrespective of the focus of attention, the global configuration of parts is processed quicker than local shape information.

Since the overall level of performance for configuration change was better than for shape change, it might be argued that the differential cueing effect reflects differences in task difficulty rather than aspects of shape processing *per se*. However, data from a pilot version of Experiment 1 ($n = 34$) indicates that this was not the case. This pilot was identical to the current experiment except that the initial stimulus duration was 2.5 s and the subsequent stimulus duration was self paced (compared to 1.5 s and 100 ms, respectively, in the current experiment). While the overall accuracy in the pilot was much higher (close to ceiling—93% correct for configuration change and 82% for shape change in no cue conditions compared to 84% and 71%, respectively, in Experiment 1),[1] the overall pattern of results (including the differential cueing effect) was similar to those of Experiment 1. The interaction between cue and change type in a mixed design ANOVA comparing performance in the pilot experiment and Experiment 1 was significant, $F(2, 104) = 6.9$, $p < .01$. This differential cueing effect was similar across experiments, as the three-way interaction between cue, change type, and experiment was not significant, $F(2, 104) = 0.9$, $p = .43$. Given that the differential cuing effect was relatively consistent across large variations in accuracy, we can rule out task difficulty as an explanation of the current findings.

[1] In fact, the accuracy in the no cue shape change condition in the pilot was not significantly different to accuracy in the no cue configuration condition in Experiment 1, $t(52) = -0.059$, $p = .56$.

EXPERIMENT 2A

Experiment 1 investigated the deployment of attention to different locations of change within a single 3-D object (i.e., attention focused on a small region of a single object). The results showed that regardless of the locus of attention within an object, configural changes are better detected than shape changes. However, attention may also be distributed over a larger space or number of items. In visual search tasks, where targets must be detected amongst distractor items, the distribution of attention can be explicitly manipulated by varying the number of distractors. The larger the number of distractors, the wider the distribution of attention needs to be in order to perform the task. Thus, if the configural advantage found for single object displays is independent of the distribution of attention, we should find that the search for targets determined by configuration is more efficient than the search for targets determined by shape. An alternative explanation for this potential result would be that changes to configural information are better detected by peripheral vision.[2] However, if the configural advantage found for single object displays is based on a narrow focus of attention within an object, then adding distractors to the display should reduce this advantage by distributing attention more widely.

Saumier et al. (2002) found that search rates were faster for targets that differed from distractors in terms of the configuration of parts (compared to when they had the same configuration) and that search was worst when targets and distractors shared both their parts and configuration. Based on these results, it was expected that search for targets involving a configuration change would be quicker than search for targets involving no configuration change. Specifically, the current experiment tested this hypothesis by investigating search for three different target types. The target differed from distractors in terms of: (a) The configuration of parts (i.e., targets and distractors have the same parts), (b) one part having a different shape (i.e., targets and distractors have the same configuration), or (c) a simple switching of parts (i.e., targets and distractors have the same configuration and same parts, but a different relative arrangement). Thus, search was expected to be quicker for configuration targets than shape or switch targets. However, there were important differences in design between the current study and that of Saumier et al. Saumier et al. used objects with familiar configurations (e.g., animals) and changes that involved every part of the object. This might have led to inflated performance (indeed, slopes were nearly flat for both the agnosic and control subjects for same configuration conditions). Conversely, the stimuli used in our study were novel objects with unfamiliar configurations and the changes made were subtler, involving only one or two parts. The use of this type of stimuli should allow us to explore perceptual processing while controlling for higher level labelling or semantic processing.

[2] Thanks to an anonymous reviewer for this suggestion.

Method

Subjects. A total of 29 undergraduate students participated and were tested individually. Subjects received course credit for participating.

Materials. Stimuli were 3-D novel objects similar to those used in Experiment 1. Each object was composed of a main body with three appendage parts. The appendages were attached to the body at three of six possible positions. A total of 12 different object exemplars were used in the current experiment. The target–distractor relationship was such that the target differed from the distractors in terms of: (a) The configuration of the object parts; (b) a switching of two of the object parts; or (c) the shape of one of the parts. All objects were photorealistically rendered with the same colour and texture. The entire background screen was white. The objects were all of similar size, with the average dimensions of each object being 55 pixels wide and 65 pixels high.

Procedure. The experiment consisted of 24 randomly ordered blocks of 30 trials, resulting in a total of 720 trials. There were eight blocks of each of the target–distractor difference types (configuration, shape, and switch). The targets and distractors for each block were counterbalanced. At the beginning of each block, subjects were shown the target and distractor (see Figure 3). Time allowed to study the instruction screen was self paced. Each trial began with a

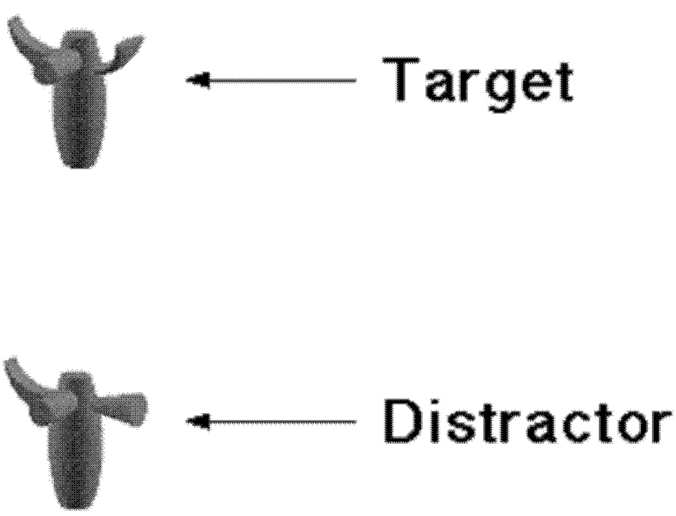

Figure 3. An example of an instruction screen indicating the target and distractor items shown at the beginning of each block in Experiment 2a.

fixation cross appearing for 0.5 s at the centre of the screen, followed by the object display. Objects remained on the screen until a response was made.

Each visual display showed 2, 6, or 10 items. Stimuli were shown at 30 possible locations (each jittered by 4 pixels) across the computer screen. The target appeared (in a random location) in half of the trials. In the remaining half of trials, only distractors were present (target absent trials). The target present trials were split equally into the three set size conditions, that is, 10 trials of each size per block. Participants were asked to indicate whether the target was present or absent by pressing corresponding keys on a keyboard. Feedback was given in the form of a beep to incorrect responses.

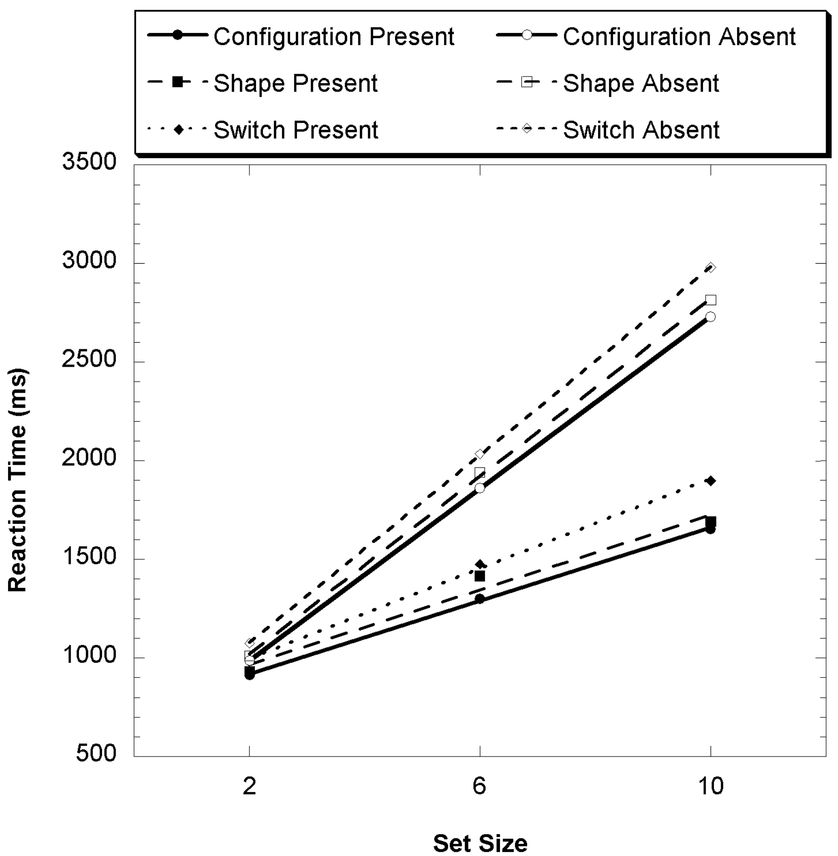

Figure 4. Mean reaction time on the visual search task as a function of target type and set size. Results are shown for both target present and target absent trials.

Results and discussion

Data analysis was conducted using accurate responses. Because there were instances of reaction times (RT) greater than 20 s, RTs of more than 2.5 standard deviations from the mean for each condition were omitted from the analysis. Figure 4 shows reaction times and slopes for the three target–distractor difference types for both target absent and present conditions. A one-way ANOVA on RT showed a significant variation among different target types, $F(2, 56) = 40.039$, $p < .01$. Bonferroni-adjusted post hoc contrasts showed that RTs for each of the target conditions were significantly different (all ps $< .05$). RT was slowest for switch targets, quicker for shape targets, and quickest for configuration targets. The interaction between condition and set size was significant, $F(4, 112) = 11.88$, $p < .01$, indicating a difference in slopes among conditions.

Regression slopes for RT by set size for each subject were calculated. A one-way ANOVA on the slopes showed a significant variation between target type, $F(2, 56) = 16.52$, $p < .01$. Bonferroni-adjusted post hoc contrasts showed that the slope for the switch target condition (114.0 ms/item) was significantly greater than the slopes for configuration (92.8 ms/item) and shape (95.0 ms/item) change targets (both ps $< .01$). The slope for the shape change was not significantly different to that found for the configuration change target condition ($p > .05$). These results indicate similar efficiency when searching for targets based on configuration and shape information, while search for targets based on switch information was less efficient.

Upon scrutiny of the slopes in Figure 4, there appeared to be differences in the target absent conditions. We conducted a one-way ANOVA on RT for target absent trials, which showed a significant variation between target type, $F(2, 56) = 40.66$, $p < .01$. The interaction between condition and set size for target absent trials was significant, $F(4, 112) = 7.80$, $p < .01$, suggesting that the difference in slopes might have been due to subjects using different strategies in each of the conditions. Given the blocked nature of the task, this was not particularly surprising. At the beginning of each block, subjects were shown an instruction screen with the target and distractor items (see Figure 3). Subjects could then focus on the part or parts of the objects that had the most useful or diagnostic information for successful change detection in that block. For example, in a block defined by the targets and distractors in Figure 3, subjects needed only to focus on whether the part on the left-hand side was curved or a cone. One problem with this potential strategy is that subjects would have been relying only on part information rather than encoding the objects as wholes. That is, subjects would have been searching for specific areas or parts of an object rather than searching for a whole object target. Since the aim of this study was to investigate the processing of whole objects, not sections of objects, a second visual search experiment was conducted in which trials were not blocked.

EXPERIMENT 2B

To investigate whether blocking in the previous experiment influenced performance, in the current experiment the same targets and distractors were used; however, in this case trials were fully randomized. No information about the target and distractors was provided to subjects; their task was simply to indicate whether the objects in the display were all the same or whether one was different (''odd man out'' task). One advantage of this task is that it can be used to explore the kinds of information being spontaneously used in visual object discrimination. Subjects were not made aware of the type of difference between the targets and distractors. Thus, if a target defined by a configuration difference were to be detected as an ''odd man out'' quicker than a target defined by a shape difference, for example, it would be presumably be due to information about configural properties being processed quicker than information about the shape properties of the object.

Method

Subjects. A total of 29 undergraduate students participated and were tested individually. Subjects received course credit for participating.

Materials. The materials and stimuli were the same as for Experiment 2a.

Procedure. Trials were fully randomized, with a total of 720 trials (the same number as Experiment 2a). An ''odd man out'' task was used in which subjects were asked to indicate whether all of the objects in the display were the same or if one object in the display was different from the rest. Each trial began with a fixation cross appearing for 500 ms at the centre of the screen, followed by the object display. Objects remained on the screen until a response was made.

Each visual display showed 2, 6 or 10 items. Stimuli were shown at 30 possible locations (each jittered by 4 pixels) on the computer screen. Half of the trials were *same* trials, the other half *different* trials. Same trial displays consisted of all the same objects. Different trial displays had one object different to the rest. The different trials were split equally into the three change type conditions. Participants were asked to indicate whether the objects were all the same or one of the objects was different by pressing corresponding ''same'' and ''different'' keys on a keyboard. Feedback was given in the form of a beep to incorrect trials.

Results and discussion

Data analysis was conducted using accurate responses. Again, RTs more than 2.5 standard deviations from the mean for each condition were omitted from the analysis. RT and slopes for the three target–distractor difference types are

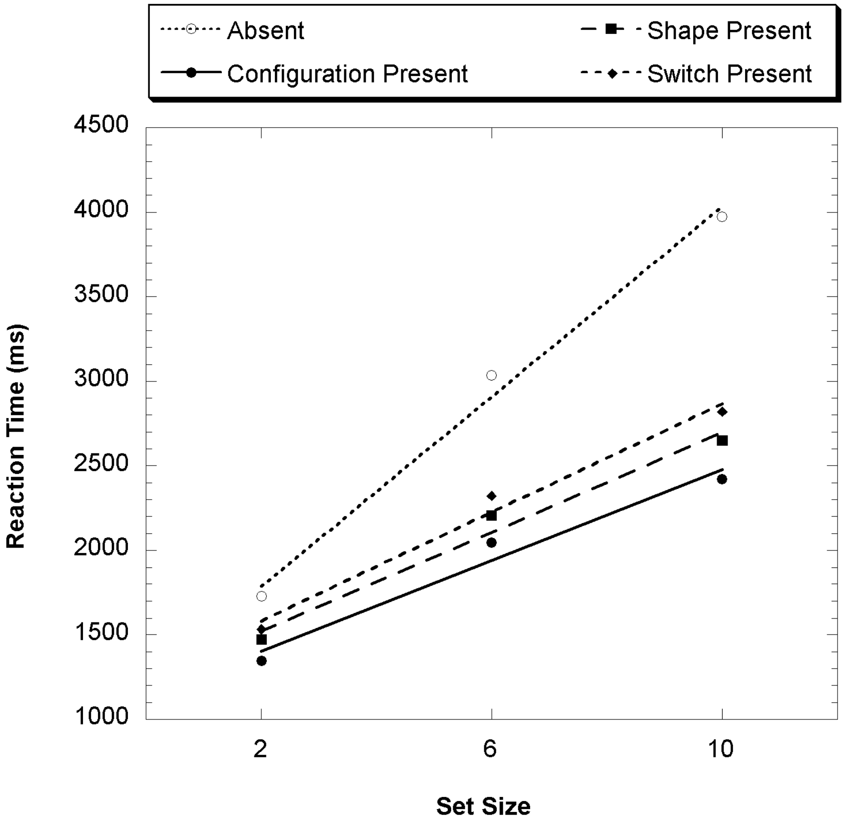

Figure 5. Mean reaction time on the ''odd man out'' visual search task as a function of target type and set size.

illustrated in Figure 5 for both target absent and present conditions. Importantly, the pattern of differences between conditions was similar to that of the previous experiment. Because an ''odd man out'' task was used, there was only one target absent condition (trials in which all objects in the display are the same). A one-way ANOVA on RT showed a significant variation among different target types, $F(2, 56) = 52.27$, $p < .01$. Bonferroni-adjusted post hoc contrasts showed that RTs for each of the target conditions were significantly different (all ps $< .01$). As shown in Figure 5, RT was slowest for switch targets, quicker for shape targets, and quickest for configuration targets. The interaction between condition and set size was significant, $F(4, 112) = 2.69$, $p < .05$, indicating a difference in slopes between conditions.

Regression slopes for RT by set size for each subject were calculated. A one-way ANOVA on the slopes showed a significant variation between target type,

$F(2, 48) = 5.22$, $p < .01$. Bonferroni-adjusted post hoc contrasts showed that the slope for the switch target condition (160.7 ms/item) was significantly higher than the slope for the configuration (134.3 ms/item) condition ($p < .01$). The shape (147.0 ms/item) change slope was not significantly different to either the configuration or switch conditions (both *p*s > .05). Slope analysis indicated that configuration and shape information was processed with similar efficiency, while switch information was less efficiently processed (although now not significantly less than shape). Overall, RT was longer than the previous experiment; this was likely to be due to the fully randomized trial design and the fact that the relationship between the targets and distractors was not made explicit in the current experiment.

The target present results showed the same pattern in both Experiments 2a and 2b. The only difference between the two experiments was that subjects were made explicitly aware of the relationship between the target and distractors in Experiment 2a, but no prior information about the target and distractors was given in Experiment 2b. Regardless of these differences, the results of both visual search experiments showed that configural differences between targets and distractors were detected faster than part shape differences or part switches. Analysis of the slope functions, however, showed that there was no significant difference in search efficiency for configural and local shape information. If it were the case that attention was more easily allocated in peripheral vision to configural differences between targets and distractors than to shape differences, an interaction should have been evident such that the slope for configural targets was significantly shallower than the slopes for shape or switch targets. Thus, attention appears not to be responsible for the configural advantage.

GENERAL DISCUSSION

Without the manipulation of attention, detection of change to the configuration of novel 3-D objects has been shown to be quicker and more accurate than changes to either the shape or arrangement of parts (Keane et al., 2003). In Experiment 1, we found that drawing attention to the locus of change facilitates accurate detection of both configural and shape changes within an object; however, drawing attention away from the locus of change is detrimental only to shape changes. In addition, an RT benefit for configural over shape change detection was found regardless of the locus of attention. That is, while the configural advantage can be further enhanced by focused attention to the location of change, it does not require focused attention. A configural advantage was also demonstrated in the visual search tasks used in Experiments 2a and 2b—with an RT benefit for configural target detection over shape targets regardless of set size. However, slow serial searches were required for both configural and shape targets. Processing was performed with similar efficiency for both types of information. Taken together, these findings suggest that: (a)

Groups of complex, novel, 3-D objects are attended in a serial fashion (even if targets and distractors have different configurations); and (b) once the object is attended, focused attention is only required if the change or the target–distractor difference is based on part shape (rather than configuration).

Rensink (2002) argued that focused attention is required to detect change. Experiment 1 investigated whether the configural advantage was a consequence of focused attention being more easily directed toward configural than part shape properties of single objects. The results suggest that: (a) Focused attention on the location of change within an object is necessary for the accurate processing of part shape information but not configural information; and (b) when the locus of attention within an object is controlled for, the configural properties of an object are processed quicker than other object shape properties. Thus, it seems that focused attention is not necessary to detect some types of change. This idea is compatible with studies by O'Regan et al. (1999, 2000), showing that allocating attention to the location of change does not inevitably result in successful change detection. Rather, more global configuration or layout aspects of a scene, which set up an overall framework for representation, might be important for successful change detection. Following along these lines, there are studies on multiple object displays that suggest that overall configural information might be fundamental to organization in visual short-term memory (Jiang, Olson, & Chun, 2000; Johnston & Pashler, 1990; Sagi & Julesz, 1985).

Aginsky and Tarr (2000) examined the cuing of change detection during scene perception, a study that appears to be relevant to the findings of Experiment 1. They found an RT advantage for cuing colour but not for the properties that influenced the configuration of the scene: object position or presence. Aginsky and Tarr argued that colour showed a cuing advantage because it was a poorly encoded property of the scene, whereas the encoding of configural properties like object position and presence are better encoded in scene representations without the need for active deployment of attention. Aginsky and Tarr went further to suggest that these results lend support to a two-stage process for scene perception in which objects relevant to scene layout are processed automatically, whereas detail properties are processed only through focused attention. For example, according to some theories of visual search (e.g., Triesman & Gelade, 1980; Wolfe, 1994; Wolfe, Cave, & Franzel, 1989), the coarse information extracted by initial processing provides a sort of map of distinct regions to which attention may be focused or guided for further analysis of detail features such as texture or complex shape. On surface inspection the current results, too, appear to be compatible with this idea of a two-stage coarse to fine processing for visual perception in which global, configural information is processed automatically and prior to any finer detail. However, the long RTs for all change types, including configural change, in both our study and in those of Agnisky and Tarr, render this explanation unlikely.

To determine whether the configural advantage persisted when the distribution of attention was widened, two visual search experiments were conducted.

Regardless of whether subjects were made explicitly aware of the relationship between the target and distractor, the pattern of performance in the two experiments was the same. The results showed that search for configural differences between targets and distractors was quicker than search for part shape differences or part switches across all set sizes, suggesting that configural information may be utilized quicker than local shape information. However, the slopes of the RT by set size functions indicated that search for a configural, shape, or switch target was far from efficient; search for all target types was slow and serial. Importantly, information about the shape of parts appeared to be processed as efficiently as configural information. Thus, it can be concluded that although configural differences were not automatically detected in a scene-like display of objects, widening the distribution of attention does not weaken the configural advantage within objects.

Taken together then, the current results are not compatible with a two-stage process in which configural information is processed automatically and more complex shape information is processed subsequently. Rather, they suggest that to detect a 3-D target object within a multiobject display (irrespective of the type of change or difference between the target and distractor) an observer needs to attend to each object in a serial fashion. Once attention is focused on an object *per se*, it is not necessary to further refine that attention to specific locations within the object to detect the presence of configural differences. However, attention to the location of change is necessary for the detection of differences in the shape of parts. Research on attention to the parts of an object suggests that these two actions occur in parallel, that is, part-based attention operates at the same time as the object is processed as a whole (Vecera et al., 2000). Thus, the current results suggest that the configural advantage is not due to an attentional advantage; processing efficiency was similar for shape and configural targets in visual search tasks. It is only once an object is attended that the processing of information regarding the configuration of parts appears to be more accurate and faster than the processing of local part shape information.

REFERENCES

Aginsky, V., & Tarr, M. J. (2000). How are different properties of a scene encoded in visual memory? *Visual Cognition*, *7*, 147–162.

Biederman, I. (1987). Recognition-by-components: A theory of human vision understanding. *Psychological Review*, *94*, 115–145.

Cave, C. B., & Kosslyn, S. M. (1993). The role of parts and spatial relations in object identification. *Perception*, *22*, 229–248.

Chun, M. M., & Wolfe, J. M. (2000). Visual attention. In E. B. Goldstein (Ed.), *Blackwell handbook of perception* (pp. 272–310). Oxford, UK: Blackwell Publishers Ltd.

Fernandez-Duque, D., & Thornton, I. M. (2000). Change detection with and without awareness: Do explicit reports underestimate the representation of change in the visual system. *Visual Cognition*, *7*, 17–42.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 683–702.

Johnston, J. C., & Pashler, H. E. (1990). Close binding of identity and location in visual feature perception. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 843–856.

Keane, S. K., Hayward, W. G., & Burke, D. (2003). Detection of three types of changes to novel objects. *Visual Cognition*, *10*, 101–127.

Kimchi, R., & Bloch, B. (1998). Dominance of configural properties in visual form perception. *Psychonomic Bulletin and Review*, *5*(1), 135–139.

Klein, R., Kingstone, A., & Pontefract, A. (1992). Orienting of visual attention. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 46–65). New York: Springer.

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin and Review*, *4*(4), 501–506.

O'Regan, J. K., Duebel, H., Clark, J. J., & Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition*, *7*, 191–211.

O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change blindness as a result of ''mud-splashes''. *Nature*, *398*, 34.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

Pomerantz, J. R. (1983). Global and local precedence: Selective attention in form and motion perception. *Journal of Experimental Psychology: General*, *112*, 516–540.

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology*, *53*, 254–277.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*, 368–373.

Sagi, D., & Julesz, B. (1985). ''Where'' and ''what'' in vision. *Science*, *228*, 1217–1219.

Saumier, D., Arguin, M., Lefebvre, C., & Lassonde, M. (2002). Visual object agnosia as a problem in integrating parts and part relations. *Brain and Cognition*, *48*, 531–537.

Scholl, B. J. (2000). Attenuated change blindness for exogenously attended items in a flicker paradigm. *Visual Cognition*, *7*, 377–396.

Simons, D. J. (1996). In sight, out of mind: When object representations fail. *Psychological Science*, *7*(5), 301–305.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*, 261–267.

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception and Psychophysics*, *49*, 83–90.

Theeuwes, J. (1992). Perceptual selectivity for colour and form. *Perception and Psychophysics*, *51*, 599–606.

Triesman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.

Vecera, S. P., Behrmann, M., & McGoldrick, J. (2000). Selective attention to the parts of an object. *Psychonomic Bulletin and Review*, *7*, 301–308.

Williams, P., & Tarr, M. J. (1998). *RSVP: Experimental control software for MacOS* [Computer software]. Retrieved October 7, 2005, from http://www.cog.brown.edu/~tarr/RSVP/

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, *1*, 202–238.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided Search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 419–433.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 601–621.

Manuscript received
Manuscript accepted